Evaluating the Federal Weatherization Assistance Program using a randomized encouragement design

Meredith Fowlie and Catherine Wolfram

UC Berkeley

Sept 11, 2009

Today's presentation

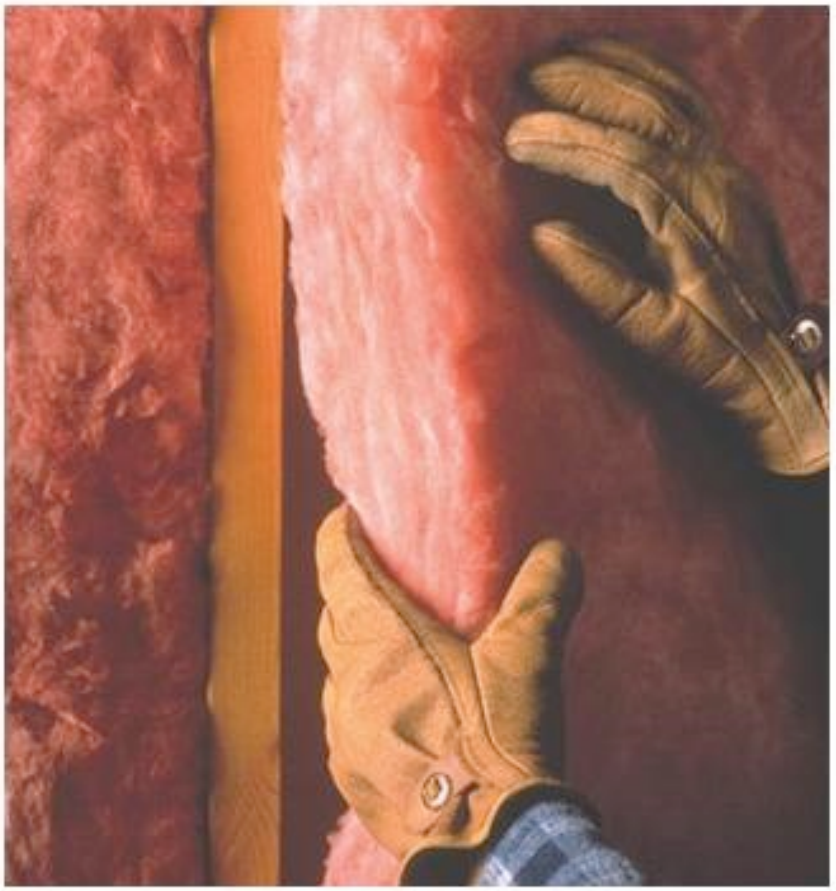
- Provide a brief introduction to the Federal Weatherization Assistance Program.
- Introduce a new field experiment- still in the design stages!
- Hoping to tap your expertise regarding some specific aspects of the research design, related research at LBL, etc.

The Weatherization Assistance Program

- Launched by President Ford in 1976.
- Over the past 30 years, an estimated 6.2 million households have received weatherization assistance.
- On the campaign trail, Obama set a goal of weatherizing one million low-income homes each year for the next decade.
- The American Recovery and Reinvestment Act allocates almost $5 billion to weatherization assistance (more than a 20-fold increase!)
- All funds must be spent or committed by Sept. 30, 2010.

Weatherization assistance

- The program serves eligible low-income households free of charge.
- “Weatherization” comprises a comprehensive series of energy efficiency measures: heating and cooling systems, its electrical system, and electricity consuming appliances
- Maximum average expenditure per household recently increased $6,500 (from $3,055).



What kind of weather are we talking about?

- Traditionally, the focus of the program has been on reducing heating costs for low income families (particularly in Northeast and Midwest).
- Increasing awareness that health and safety impacts on of extended heat waves on low income households/elderly can also be severe.
- Thanks to the stimulus, campaign to bring a better regional balance to the allocation of weatherization funds will be realized this year .

Who is eligible?

- ARRA increases eligibility criteria to 200 percent of the poverty line.
- Low-income housing may not qualify if dwelling in need of major renovation.
- If landlords do not agree to keep rent at their current level following weatherization, renters cannot participate.
- An estimated 40M hh are eligible.

(Almost) everyone loves weatherization!

- Improve the living conditions of poor households.
- Decreases dependence on foreign sources of energy.
- Harvests “lowest hanging fruit” with respect to climate change mitigation options.
- Green jobs!

"We're creating jobs immediately by … weatherizing 2 million American's homes, as was called for in the package. So that right there creates economic stimulus.… In the case of homeowners, they will see more money in their pockets, and we're reducing our dependence on foreign oil in the Middle East. Why wouldn't we want to make that kind of investment?"

President Barack Obama, Press Conference, February 9, 2009

"… if you look at the over $500 billion worth of spending … —and $6 billion to community action programs to do weatherization programs. It's just more of the same kind of wasteful spending that we have seen in the past. I was really—I was shocked."

House Minority Leader John Boehner, PBS NewsHour, January 15, 2009

Measuring WAP impacts

- In order to estimate the causal effect of weatherization assistance on energy consumption among participating households, we need credible, unbiased estimates of what energy consumption patterns would have been in the absence of the intervention.

- PROBLEM: We cannot observe this “counterfactual” outcome for weatherized households.

A more formal conceptualization of the problem:

- Let $D_i = 1$ if household i receives weatherization.
- Let $D_i = 0$ if household i does not receive weatherization.
- Let Y_i represent the outcome of interest (such as energy consumption or change in energy consumption).
- Each household is associated with two potential outcomes: $Yi(1)$ and $Y_i(0)$.
- In any given time period, we can only observe one outcome per household.
- Estimand of interest: average treatment effect on the treated: $ATT = E[Y_i(1)|D_i=1] - E[Y_i(0)|D_i=1]$

Previous National WAP evaluations

- DOE sponsored a comprehensive evaluation in 1990.
- Randomly selected 400 agencies and requested data from homes weatherized in PY 1989 and awaiting weatherization (N = 30,543 single family homes)
- Successfully obtained billing data for 32% of sample.
- Average (normalized) energy savings attributable to WAP intervention calculated as:

$$\left(\overline{Y_i(1) \mid D_i = 1}\right) - \left(\overline{Y_i(0) \mid D_i = 0}\right)$$

(where Y_i is defined to be change in energy consumption at household i between 1988 and 1989).

National WAP evaluation : Present

- National evaluation will be carried out by ORNL, Apprise, and other sub-contractors (funded under ARRA).
- Evaluation will estimate energy savings and cost effectiveness for Program Years 2007 and 2008.
- Research design very similar to 1990 evaluation.
- Official Start-date: August 1, 2009.
- Innovations in the delivery of weatherization services facilitated with ARRA money will probably be evaluated as well.

Key underlying assumptions

- Energy consumption patterns observed in the control group (*i.e.* $Y_i(0)|D_i=0$) are equal, in expectation, to the consumption patterns that would have been observed at the weatherized households (*i.e.* $Y_i(0)|D_i=1$) absent the treatment.

- Any observable differences between the treated and control groups (i.e. . $Y_i(1)|D_i=1 - Y_i(0)|D_i=0$) are either purely by chance or directly caused by the intervention.

- Power calculations and econometric methods are predicated upon these assumptions.

Pros and cons of retrospective analyses

Retrospective studies have many advantages:

- Relatively inexpensive.
- Comprehensive: Can be used to analyze WAP across the country.
- Non-disruptive- no changes to program implementation required. Data can be collected ex post.

- But, there is no guarantee they will yield unbiased estimates of the effects of WAP.

Possible biases in retrospective analyses

Suppose households are more likely to apply for WAP after they've experienced a positive shock to consumption (here we assume average consumption 100 w/o weatherization) :

	2007	2008
Y (1) | D=1 (observed)	110	80
Y (0) | D=1 (counterfactual)	110	100
Control : Y(0)|D=0 (observed)	100	110

- Retrospective studies will calculate a weatherization effect of : (110-80) – (100-110) = 40.
- True effect : 100-80 = 20.

Other potential concerns

- Treatment timing is determined in part by CAAs by factors that can also affect outcomes (including Federally mandated prioritization protocols).
- Other factors (such changes in outreach activities, economic conditions, etc.) could lead to non-random differences across groups.

Punchline : Impossible to know if consumption among treatment and controls would have been equal (in expectation) absent the intervention.

Observational studies in a larger context

- Absent random assignment, there is always greater risk that systematic differences might be responsible for some/all of the observed differences in the outcome of interest.
- There have been several within study comparisons of experiments and non-experiments to assess the internal validity of retrospective, non-experimental program evaluations (*e.g.* Bloom, 2002; Cook *et al.,* 2006; Glazerman *et al.,* 2003).

"The answer to the question, 'Do the best observational methods work well enough to replace random assignment?' is probably, 'No.'" - *Bloom (2002)*

Randomized Controlled Trials (RCTs)

- The “gold standard” for making a causal inference about an intervention’s impact on an outcome of interest.
- RCTs eliminate (or greatly mitigate) selection bias by design; the credibility of estimates is significantly improved.
- A growing literature documents practical experience with designing and implementing RCTs in a variety of contexts .

Randomized Controlled Trials (RCTs)

Used to address (and in several instances, validate) a broad range of domestic policies and programs:

- Welfare to work programs (Hamilton, 2002)
- Policing strategies (Sherman and Weisburd, 1995)
- Public education programs (Kemple and Snipes, 2000)
- Housing assistance (Orr et al, 2003)
- Health insurance programs (Newhouse 1996),
- Job training programs (Bloom et al, 1997),
- Unemployment insurance programs (Robins and Spergelman, 2001)
- etc.

Standard RCT design

- Individuals are randomly drawn from the population of interest.
- This sample is randomly divided across intervention (*i.e.* treatment) group and a control group; two groups are identical in expectation by design.
- Post-intervention, outcomes are compared across groups .
- PROBLEM : Mandating participation of some while preventing participation of others is impossible here.

A randomized encouragement design

REDs are particularly useful when:

- Randomization of access or mandatory participation is not practical or desirable.
- No need to ration available services (i.e. demand does not exceed supply).
- The effects of both participation and outreach are of policy interest.

Rather than randomize over the intervention itself, we randomly manipulate encouragement to participate.

In theory, households in our sample fall into three categories

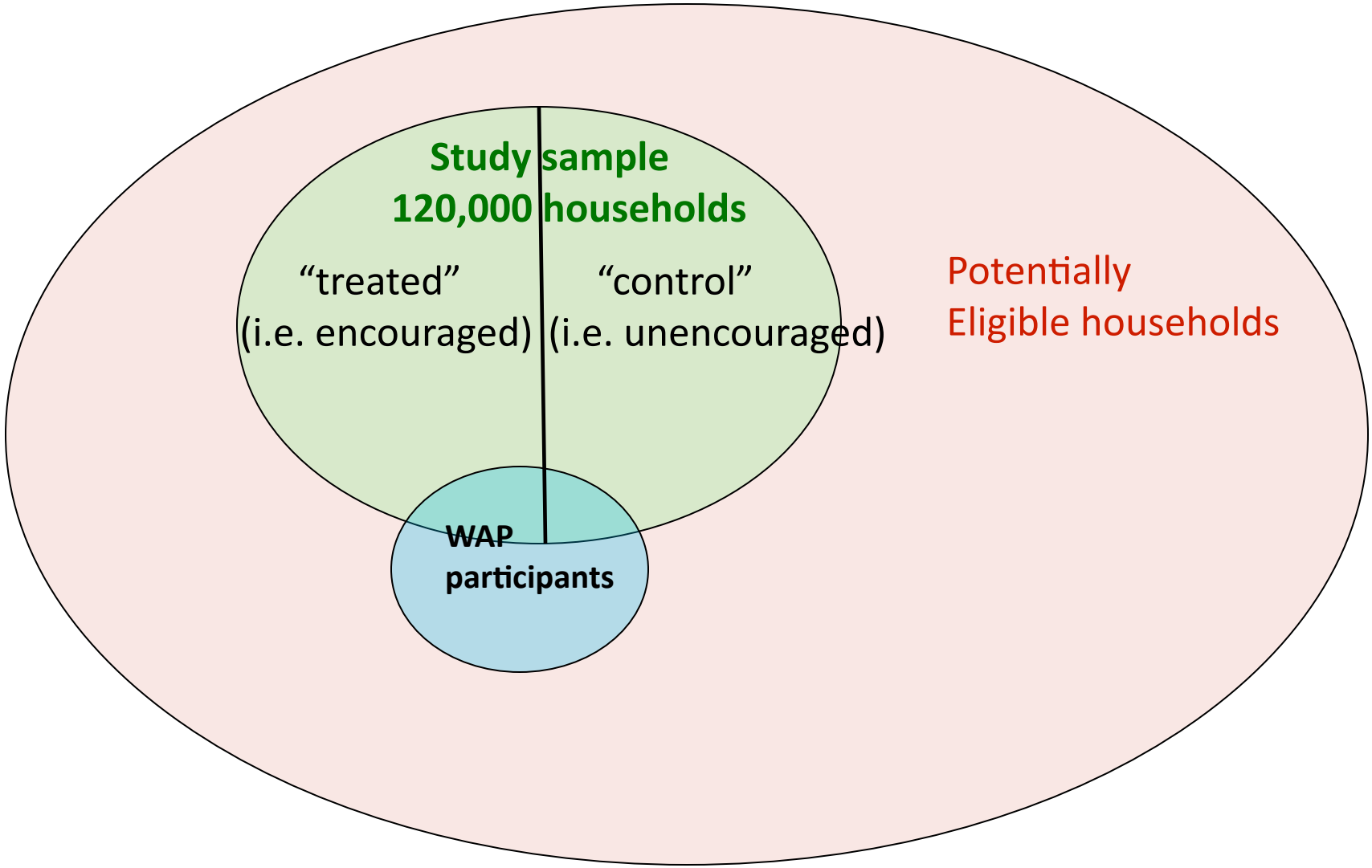
"**Never takers**" : Households who will not seek out weatherization assistance, regardless of whether they receive encouragement/ outreach information.

"**Always takers**" : Households who will learn about the program and seek out assistance with or without our encouragement.

"**Compliers**" : Households who we will persuade with our encouragement to participate, but who would not otherwise have received weatherization assistance.

(This assumes that our encouragement instrument has a weakly positive effect on program participation or all households)

Our randomized encouragement design (the cartoon version…)



Necessary ingredients

- Close cooperation with implementing agencies.
- WAP data (identifying weatherized households, which interventions are installed, timing of treatment, etc).
- Household billing data (electricity and natural gas) for all 120,000 households in the study.
- An effective encouragement instrument !

(Tentative) timeline

- Work on the design of the encouragement instrument , sampling frame : Fall/winter 2009.
- Stimulus to WAP expected to ramp up over PY2009, taking full effect in PY 2010.
- We hope to be in the field with encouragement instrument by early 2010.
- Household billing data for 2008-2011/12 will be used, together with WAP data.

Primary research questions

- What effect does weatherization assistance have on energy consumption/expenditures at participating households?
- Do estimated treatment effects vary systematically with observable socio-economic characteristics?
- Do we find evidence consistent with a rebound effect?

Estimating direct energy impacts

Fairly straightforward given our research design:

$$Y_{it} = \alpha_i + \pi z_{it} + \beta' X_{it} + \varepsilon_{it}$$

α_i : household fixed effect

z_i : indicates encouragement

X_i : observable determinants of energy consumption

ε_{it} : unobservable (by us) determinants of energy consumption.

- Extrapolating from compliers to larger treated population will require additional assumptions.

Exploring treatment effect heterogeneity

- Use census block data to explore variation across socio-economic groups.
- Utility partners have collected some data on appliance penetration/dwelling characteristics.
- Possibly implement an occupant survey to collect dwelling characteristics for a subset of treatment and control groups.
- Other available sources of data on dwelling characteristics?

Estimating possible rebound effects

- Small/zero impact on energy consumption need not imply a failed program!
- We are considering various approaches to investigating rebound effects:
  - Use simulation-based models to put bounds on the rebound effect.
  - Self-reported measures of comfort level/ thermostat settings.
  - In-home temperature data loggers?

Other areas we would like to explore..

- Non-energy benefits including:
  - Interactions with LIHEAP
  - Health-related outcomes
- Relative effectiveness of alternative persuasion/motivation strategies among different subsets of eligible population.

All suggestions/feedback welcome!

Meredith Fowlie : fowlie@berkeley.edu

Catherine Wolfram: wolfram@haas.berkeley.edu

Thank you!!